

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-339104
(43)Date of publication of application : 08.12.2000

(51)Int.Cl.
G06F 3/06
G06F 12/08
G06F 12/16

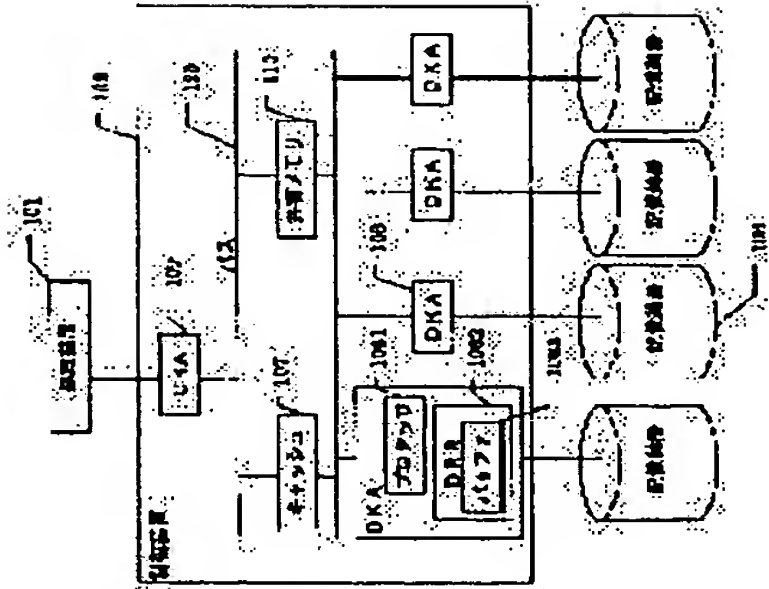
(21)Application number : 2000-081711 (71)Applicant : HITACHI LTD
(22)Date of filing : 17.03.2000 (72)Inventor : ARAI HIROHARU
SUZUKI SUSUMU
YASUKAWA HIRONORI

(30)Priority
Priority number : 11075174 Priority date : 19.03.1999 Priority country : JP

(54) DISK SUB-SYSTEM AND DATA COPYING METHOD

(57)Abstract:
PROBLEM TO BE SOLVED: To generate the copy of a logic volume without a host channel by changing a logic address of data on a regular volume which is read into the logic address of an auxiliary volume and transferring it to the auxiliary volume.

SOLUTION: A controller 102 gives the instruction of a copy of data on a cache 107 to a hardware of a DRR 1082. The controller 102 secures an area for placing the copy on the cache 107. A processor 1081 of a DKA 108 reads objective data (data of the regular volume) which is read to the cache 107 to a buffer 108 in the DRR 1082. The regular address of data which is read is changed to a volume number which is set when a new pair generation command is issued by using the DRR 1082. A check code is also changed by adjusting it to the changed volume number. Data whose logic address is changed is transferred to an area for auxiliary volume in the cache 107.



LEGAL STATUS

[Date of request for examination]
[Date of sending the examiner's decision of rejection]
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]
[Date of final disposal for application]
[Patent number]

BEST AVAILABLE COPY

[Date of registration]
[Number of appeal against examiner's decision of rejection]
[Date of requesting appeal against examiner's decision of rejection]
[Date of extinction of right]

Copyright (C): 1998,2003 Japan Patent Office

★ NOTICES ★

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] The disk subsystem equipped with the storage which records data, the disk adapter which controls the data transfer to this storage, a processor and the channel adapter which performs data processing, the cache memory which records temporarily the data from the aforementioned storage and/or the aforementioned processor, and a means to change the distinction information on the data recorded on this cache memory.

[Claim 2] The aforementioned distinction information is a disk subsystem according to claim 1 which is the logical address.

[Claim 3] The disk subsystem according to claim 1 which uses DRR which creates the parity of the data recorded on the aforementioned storage as a means to change the aforementioned distinction information.

[Claim 4] The disk subsystem characterized by providing the following. A buffer means to be connected with a storage means and/or a processing means, and to record the data from these temporarily. A means to control the data transfer between this buffer means and the aforementioned processing means. A means to control the data transfer between the aforementioned buffer means and the aforementioned storage means. A means to specify the logical volume of a couple as right volume and subvolume within the aforementioned storage means, the means which reads the data of the aforementioned positive volume to the aforementioned buffer means, a means to change the logical address of the read aforementioned data into the logical address of the aforementioned subvolume, and a means to transmit the data which changed the aforementioned logical address to the aforementioned subvolume.

[Claim 5] The copy method of the data recorded on storage characterized by providing the following. The step which receives the command which copies data from a processor. The step which sets up the distinction information on the copy place volume of the aforementioned data. The step which sets up the field which records copy data in a cache memory. the step which reads the aforementioned data from a cache memory, the step which changes into the distinction information on the aforementioned copy place the logical address of the data which carried out [aforementioned] reading appearance, and the step which sends the data which changed the logical address to the field which carried out [aforementioned] a setup

[Claim 6] The aforementioned distinction information is the data copy method according to claim 5 which is the logical address.

[Claim 7] The data copy method according to claim 5 of reading the aforementioned data from a cache memory to the buffer in DRR which creates the parity of the data recorded on storage, and changing the aforementioned distinction information with Above DRR.

[Translation done.]

* NOTICES *

Japan Patent Office is not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.*** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[The technical field to which invention belongs] The invention in this application relates to the pair of the logical volume within the same control unit, and duplicate creation in the storage system which consists of two or more logical volumes.

[0002]

[Description of the Prior Art] In case the duplicate of a logical volume is created, the data which serve as a pair first are created and a duplicate is created by dividing this pair after that. If simple copy operation is performed in this pair creation stage, even the logical address and the check code which are used as distinction information on data will be the same not to mention data. Then, in order to set up the new storing place of a pair, you have to set up the logical address newly. Moreover, in case the duplicate which subsequent I/O was based and was created at once is updated, you have to update a check code.

[0003] in order to perform these new thing setup, after creating the pair of a logical volume, the duplicate was created by using the so-called remote copy function like a publication for a ***** No. 509565 [eight to] official report as conventional technology, acting as if the control unit was a processor, receiving this data by the send and other host channels (congenital adrenal hyperplasia) to a host interface, and outputting the data of a logical volume to a logical volume That is, the data of the logical volume which creates a duplicate were sent out to the host interface at once, by receiving by the host channel as new data, the logical address of a pair was performed and a change of a check code was made.

[0004] However, with the conventional technology, in order that a control unit may act as a processor, a host channel is used. In duplicate creation within the same control unit, two or more host channels are used, and the part and the number of channels which can be used for the usual host connection decrease.

[0005]

[Problem(s) to be Solved by the Invention] When the duplicate of a logical volume is created using the above-mentioned remote copy function, in order to transmit duplicate data, the trouble that it cannot be used for connection with a processor generates these host channels that use a host channel.

[0006] The purpose of the invention in this application is to create the duplicate of a logical volume, without using a host channel.

[0007]

[Means for Solving the Problem] A buffer means for the above-mentioned technical problem to be connected with a storage means and/or a processing means, and to record the data from these temporarily, A means to control the data transfer between this buffer means and a processing means, A means to control the data transfer between a buffer means and a storage means, and a means to specify the logical volume of a couple as right volume and subvolume within a storage means, The means which reads the data of right volume to a buffer means, and a means to change the logical address of subvolume for the logical address of the read data, It can attain by considering as a disk subsystem equipped with a means to transmit the data which

changed the logical address to subvolume.

[0008]

[Embodiments of the Invention] Drawing 1 shows the system configuration of this example. The cache memory 107 as a buffer means by which a control unit 102 saves temporarily the data from a processor 101, and the data from storage 103. Data transfer with a processor 101 It consists of buses 120 which connect these with a certain channel adapter (it is indicated as congenital adrenal hyperplasia drawing) 109 to control, the disk adapter (it is indicated as DKA drawing) 108 which controls data transfer with storage 101, and the shared memory 110 which records the control information of a control unit 102. The storage 103 equipped with storages, such as a magnetic disk unit, is controlled, and the demand from a processor 101 is performed. [0009] The control information recorded on the shared memory 110 is transmitted to a channel adapter 109 or a disk adapter 108 via a bus from a shared memory 110, and is used by the processor in a channel adapter 109 and a disk adapter 108. In addition, in the invention in this application, it becomes data from the logical address and the check code of user data which are used as distinction information on user data and user data. Furthermore, a check code consists of data which check user data and the logical address.

[0010] congenital adrenal hyperplasia109 and DKA108 equip the interior with the bus adapter (not shown) which is the hardware linked to a bus 120, the processor 1081 (not shown [the inside of congenital adrenal hyperplasia109]) which controls a transfer for data. Furthermore, DKA108 is equipped with DRR (Data Recover and Reconstruct circuit)1082 which performs parity creation of the data recorded on storage 103 storage 103, and broken recovery of data according to directions of a processor 1081. Furthermore, it has the buffer 1083 used in case parity creation and data recovery are performed in the DRR1082 interior. In addition, this buffer 1083 does not necessarily exist in the DRR1082 interior.

[0011] First, the commands set about a pair, and creation and deletion of a duplicate is shown in drawing 2 . A commands set consists of the resynchronization command 214 again changed into a duplicate sheep state with a pair after the pair creation command 211 which specifies the right volume and the subvolume for duplicate creation, and creates a pair, the pair delete command 212 which cancels the pair of right volume and subvolume, the command 213 which creates a duplicate in subvolume, and duplicate creation.

[0012] Next, control information is explained. The detail of the pair information 300 on a copied material logical volume (right volume) and a copy place logical volume (subvolume) is shown in drawing 3 . It has the information on the duplicate pointer 305 which shows how far the duplicate of the pair number 301, the pair state 302, the right volume number (right volume logical address) 303, the subvolume number (subvolume logical address) 304, and right volume has ended the pair information 300 for every pair. This information is recorded on a shared memory 110 as information for the control at the time of making a pair state change.

[0013] Next, the content of the pair state 302 of the pair information 300 is explained using drawing 8 . The state transition of a pair is shown in drawing 8 . In the pair-less state 801, it is in the state which does not form the pair, and although a pair exists, it is [state / duplicate sheep / 802 / with a pair] in the state which is not generating the duplicate in subvolume. [0014] In the state where this duplicate is not generated, although all original and copy volume copies were completed, the output to subvolume is not carried out synchronizing with the I/O from a processor 101, for this reason -- periodical -- the difference from right volume to subvolume -- it is in the state where the copy of data is carried out

[0015] It is in a state, i.e., the state where the data of right volume are copied to subvolume, while generating a duplicate in subvolume, in a state 803 during a duplicate with a pair. The light to the right volume in this state is not reflected in subvolume. Therefore, if the light to right volume is generated, the content of right volume and subvolume will become inharmonious. That is, the duplicate of right volume when the duplicate creation command 213 is published is created by subvolume.

[0016] It is in the state which completed generation of a duplicate in subvolume in the state 804 with a duplicate with a pair. In this stage, all of read/write of subvolume can become possible, and it can work independently with right volume.

[0017] the difference which shows a light position to drawing 4 when a light is in right volume after the duplicate creation command 213 was published -- it records as information and the resynchronization command 214 is published, and it is used in case the duplicate sheep state 802 with a pair is created

[0018] It changes in the pair-less state 801 from every state by the pair delete command 212.
[0019] With the pair creation command 211, it changes in the duplicate sheep state 802 with a pair from the pair-less state 801. It changes in the state 803 during a duplicate with a pair from the duplicate sheep state 802 with a pair with the duplicate creation command 213. After the whole duplicate processing 500 is completed, it changes in the state 804 with a duplicate with a pair from a state 803 during a duplicate with a pair. It changes in the duplicate sheep state 802 with a pair from the state 804 with a duplicate with a pair with the resynchronization command 214.

[0020] drawing 4 -- difference -- information 401 is shown difference -- the array whose information 401 makes the pair number 301 an argument -- becoming -- *** -- each entry -- difference -- it consists of a set of a bit difference -- a bit sets up 0, when there is no difference between right volume and subvolume about corresponding data, and when a difference exists, it sets up 1 This information is recorded on a shared memory 110 as information for control. this difference -- once information 401 creates a duplicate, in case it takes adjustment with right volume and subvolume again (resynchronization is carried out), it is used
[0021] this difference -- the difference which corresponds when information 401 has the light demand of data from a processor 101 -- a bit is set as 1 moreover, the difference which corresponds when data are in agreement by carrying out duplicate processing -- a bit is set as 0

[0022] The structure of a command block is shown in drawing 7. The command block 700 consists of command code 701, the right volume number 702, and the subvolume number 703.

[0023] A control unit 102 will search an entry with the right volume number 303 and the subvolume number 304 which are in agreement with the right volume number 702 of the command block 700, and the subvolume number 703 from the pair information 300, if the command block 700 is received from a processor 101. The pair state 302 is acquired from this entry, and processing shown in drawing 2 from the command 701 of this pair state 302 and a command block is performed.

[0024] When the state of the pair specified by the pair creation command 211 is 221 without a pair, new pair creation 231 is carried out. In the new pair creation 231, it is vacant in the volume first specified from the processor 101 based on the pair information 300, an entry is found, and it assigns as area for subvolumes. And the right volume number 303 and the subvolume number 304 which were specified by the pair creation command 211 based on this assignment are set as the area of the right volume number 303 of the pair information 300, and the area of a subvolume number, respectively, the duplicate pointer 305 is initialized, and the whole duplicate processing 500 is called.

[0025] An access property which is different to the same data by specifying the logical volume from which RAID level differs as area of subvolume in the case of this new pair creation can be given. In addition, in the stage of this new pair creation, the area for subvolumes is only secured, and a copy of data is not yet performed but is [that the storing place of subvolume is only decided, and].

[0026] When the pair state when receiving the pair creation command 211 is except pair-less 221, no control units 102 are performed in principle. Although there is a pair, the case where right volume defines the logical volume from which plurality differs as subvolume, defines it as a pair different, respectively, and creates a duplicate creates a pair again. When defining it as a different pair and creating a duplicate, time which received the duplicate directive command from the processor 101 is made into the appointed time, and the case where he wants to create the history of right volume etc. can be considered by creating a duplicate one after another for every fixed time.

[0027] Moreover, when the pair state when receiving the pair delete command 212 is 221 without a pair, no control units 102 are performed. In this case, since the pair does not exist, it rejects

this pair delete command 212.

[0028] the pair state when receiving the pair delete command 212 -- a pair -- an owner -- when it is 222,223, a control unit 102 initializes the entry of the pair considered as correspondence of the pair information 300

[0029] When the pair state when receiving the duplicate formation command 213 is 221 without a pair, no control units 102 are performed.

[0030] the case where the pair state when receiving the duplicate formation command 213 is the pair owner duplicate sheep 222 -- difference -- the difference corresponding to all volume fields, for information 401 -- a bit is set as 1, the duplicate pointer 304 is initialized, and the whole duplicate processing 500 is called

[0031] the pair state when receiving the duplicate formation command 213 -- a pair owner duplicate -- an owner -- when it is 223, no control units 102 are performed

[0032] When the pair state when receiving the resynchronization command 214 is 221 without a pair, no control units 102 are performed. In this case, since the pair does not exist, it rejects this pair delete command 212.

[0033] When the pair state when receiving the resynchronization command 214 is the duplicate sheep 222 with a pair, no control units 102 are performed.

[0034] When the pair state when receiving the resynchronization command 214 is 223 with a duplicate with a pair, a pair state is changed to the duplicate sheep 222 with a pair.

[0035] Next, duplicate processing is explained in detail using drawing 5 and drawing 6.

[0036] First, the whole duplicate processing 500 is called in the pair creation command 211 or the duplicate formation command 213 at the time of receipt. The input at the time of a call specifies a pair number. When the data which want to reproduce do not exist on a cache 107, it reads from storage 103.

[0037] The entry of the pair number which is in agreement with the pair number of input is first searched with Step 501 from the pair information 300, the pair state 300 of the entry concerned is acquired, and it checks that the pair state 302 is in the duplicate sheep state 802 with a pair. If it is not in the duplicate sheep state 802 with a pair, it will judge that duplicate processing cannot be performed and processing will be ended.

[0038] 502 which will call a pair number for the unit duplicate processing 600 as input if it is in the duplicate sheep state 802 with a pair.

[0039] The detail of the unit duplicate processing 600 is shown in drawing 6.

[0040] If the unit duplicate processing 600 is called, the entry of the pair number which is in agreement with the pair number of input in Step 601 will be searched from the pair information 300, and the duplicate pointer 305 of the entry concerned will be acquired. next, the difference from the address which the duplicate pointer 305 acquired at Step 601 in Step 602 shows -- the data with which the bit is 1 -- difference -- it searches from information 401 At Step 603, the data searched with Step 602 are read from right volume to a cache 107. At Step 604, the logical volume number which acquires another field into a cache and is given to a copy of data and data in the data read at Step 603 is changed into the subvolume number 304 from the right volume number 303. At Step 605, the data copied at Step 604 are outputted to subvolume.

[0041] the difference for the data outputted to subvolume at Step 609 -- information 401 is set as 0

[0042] At Step 606, the duplicate pointer 305 is carried forward by the data outputted to subvolume. At Step 607, the duplicate pointer 305 changed at Step 606 is recorded on the pair information 300, unit duplicate processing is ended, and it takes return 608 to a calling agency.
[0043] A pair is created by the pair creation command 211 from the pair-less state 801 here, and the steps 603 and 604 when the whole duplicate processing 500 is called and the unit duplicate processing 600 is called further are explained further.

[0044] As stated also in advance, in case a pair is created first and data are copied, it is necessary to newly set up the logical address. Moreover, change of the check code accompanying a setup or change of this logical address is also needed. In the invention in this application, a setup of this new logical address is performed using DRR1082 in DKA108. Originally DRR1082 is in the disk subsystem which takes RAID composition for recovering data to creation,

parity, and other obstacle data for parity from data.

[0045] By the invention in this application, the logical address of the duplicate created in subvolume is set up using the function which assigns the logical address to parity in the case of this parity creation. For this reason, duplicate procedure of the logical volume in the invention in this application is performed by a cache 107 and DKA108. In addition, if it has separately the component equipped with the function and buffer which change the logical address in carrying out reproduction in a cache 107, it will come out not to mention the ability to copy in elements other than DKA108, and combination with a place

[0046] Step 604 is explained using drawing 10. In Step 1000, a control unit 102 takes out directions for the duplicate of the data of a cache 107 to the hardware of DRR1082. And in Step 1001, a control unit 102 secures the field for placing a duplicate on a cache 107. Next, in Step 1002, the processor 1081 of DKA108 reads the object data (data of = positive volume) read to the cache 107 in Step 603 to the buffer 1083 of the DRR1082 interior.

[0047] In Step 1003, the logical address of the data read by using DRR1082 is changed into the volume number set up when new pair creation 231 was published. It combines with the volume number changed on that occasion, and a check code is also changed. Finally the data which changed this logical address in Step 1004 are transmitted to the area for subvolumes in the cache 107 too set up by the new pair creation 231. The above performs the copy on a cache.

[0048] In addition, usually, in case a cache 107 manages data, it memorizes the logical address of data etc. to the directory, and it uses the address recorded on the directory in the case of a data transfer. Therefore, even if it changes the logical address of data on a cache 107, unless the address of this directory is changed, the output to subvolume cannot be performed.

Therefore, in this example, the copy on a cache 107 was created using the buffer 1083 in DRR1082 as mentioned above. However, after changing the data on a cache 107, and the logical address of a directory, without creating a copy on a cache 107 with the logical address of data using the buffer 1083 in DRR1082 if it has a means to also change the address of a cache directory, you may create a duplicate with outputting the data to subvolume.

[0049] The completion judging 503 of a duplicate is performed after a return from the unit duplicate processing 600. That is, when the duplicate pointer 305 is not in agreement with the tail address of a logical volume, it judges with duplicate un-completing, and it repeats from Step 501. On the other hand, when in agreement, the completion of a duplicate is judged, and it progresses to Step 504.

[0050] At Step 504, a copy in the subvolume of the data of the right volume in the time of a copy being directed completes the data which carried out the copy end by being transmitted by subvolume from a cache 107. And the pair state 302 is changed into those with a duplicate, and duplicate processing is ended.

[0051] Next, processing when the change request to the data of right volume suits a control unit 102 from a processor 101 during duplicate processing is explained using drawing 9. The flow chart of processing (precedence copy) when the change request to the data of right volume is during duplicate processing is shown in drawing 9.

[0052] The entry of the pair information 300 with the same right volume number as the volume number of input is searched with Step 901, and it asks for a pair number.

[0053] The pair state 302 is searched for from the entry of the pair number for which it asked at Step 901 by Step 902, and it judges whether it is in a state 803 during a duplicate with a pair, in not being in a state 803 during a duplicate with a pair, it performs the usual output processing at Step 909, and in being in a state 803 during a duplicate with a pair, it progresses to Step 903.

[0054] the difference from the pair number for which it asked at Step 901 in Step 903 -- the data which asked for the entry to which information 400 corresponds, and had the change request from the processor 101 -- already -- difference -- a ***** [that a bit is 1] -- judging -- difference -- output processing usual with Step 909 when a bit is not 1 -- performing -- difference -- when a bit is 1, it progresses to Step 904

[0055] At Step 904, the data which had the change request from the processor 101 are read from right volume to a cache 107.

[0056] At Step 905, the data for subvolumes are reproduced on a cache 107 from the data read

on the cache 107 at Step 904 (it is the same as Step 604 of drawing 6 shown previously).
[0057] At Step 906, the data reproduced at Step 905 are outputted to subvolume.

[0058] the difference of the data outputted to subvolume at the above-mentioned step in Step 907 -- it is set as 0 from a bit

[0059] At Step 909, the data which had the change request from the processor 101 are received from a processor 101, and the usual output processing outputted to right volume is performed.

[0060] As mentioned above, if the invention in this application specifies the volume used as right volume and a pair and creation of ** A is directed, the duplicate of right volume will be created as subvolume. In this state (before the duplicate creation command 213 in drawing 2 is

published), the writing from a processor to subvolume is forbidden. If division of a pair is directed, the duplicate creation command 213 will be executed and employment of the volume of each original and copy will be independently attained from a host the difference which the difference of right volume and subvolume shows to drawing 4 while the resynchronization command 214 will be executed and the writing to subvolume will be forbidden, if resynchronization is directed -- it is copied to subvolume based on information, and again, subvolume serves as a duplicate of right volume, and will be in the state which can be divided again

[0061] Moreover, as explanation was given [above-mentioned], at the invention in this

application, the duplicate of right volume can be created in another subvolume in the same disk subsystem by using DRR inside DKA at the time of duplicate creation of right volume, without using congenital adrenal hyperplasia.

[0062]

[Effect of the Invention] The duplicate of the specified logical volume can be created by this invention, without using the host channel slack congenital adrenal hyperplasia. Thereby, curtailment of the load of the control unit accompanying duplicate creation can be aimed at.

[Translation done.]

* NOTICES *

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.*** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]
 [Drawing 1] It is drawing showing a system configuration.
 [Drawing 2] It is drawing showing the command activity of drawing 1 .
 [Drawing 3] It is drawing showing pair information.
 [Drawing 4] difference -- it is drawing showing information
 [Drawing 5] It is the flow chart which shows duplicate processing.
 [Drawing 6] It is the flow chart which shows the order of unit duplicate processing.
 [Drawing 7] It is drawing showing a command block format.
 [Drawing 8] It is drawing showing a pair state transition.
 [Drawing 9] It is the flow chart which shows the order of processing when a data change request occurs during duplicate processing.
 [Drawing 10] It is the flow chart which shows distinction information change processing.
 [Description of Notations]
 101 [-- Storage, 107 / -- A cache memory, 108 / -- A disk adapter, 109 / -- A channel adapter, 110 / -- A shared memory, 1081 / -- A processor, 1082 / -- DRR, 1083 / -- Buffer.]
 -- A processor, 102 -- A control unit, 103

[Translation done.]

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号
特開2000-339104
(P2000-339104A)

(43)公開日 平成12年12月8日(2000.12.8)

(51)Int.Cl. ⁷	識別記号	F I	テーマト*(参考)
G 0 6 F 3/06	3 0 4	G 0 6 F 3/06	3 0 4 F
12/08	3 2 0	12/08	- 3 2 0
12/16	3 1 0	12/16	3 1 0 M

審査請求 未請求 請求項の数7 O L (全 11 頁)

(21)出願番号	特願2000-81711(P2000-81711)	(71)出願人	000005108 株式会社日立製作所 東京都千代田区神田駿河台四丁目6番地
(22)出願日	平成12年3月17日(2000.3.17)	(72)発明者	荒井 弘治 神奈川県小田原市国府津2880番地 株式会 社日立製作所ストレージシステム事業部内
(31)優先権主張番号	特願平11-75174	(72)発明者	鈴木 進 神奈川県小田原市国府津2880番地 株式会 社日立製作所ストレージシステム事業部内
(32)優先日	平成11年3月19日(1999.3.19)	(72)発明者	安川 博則 神奈川県小田原市国府津2880番地 株式会 社日立製作所ストレージシステム事業部内
(33)優先権主張国	日本(J P)	(74)代理人	100075096 弁理士 作田 康夫

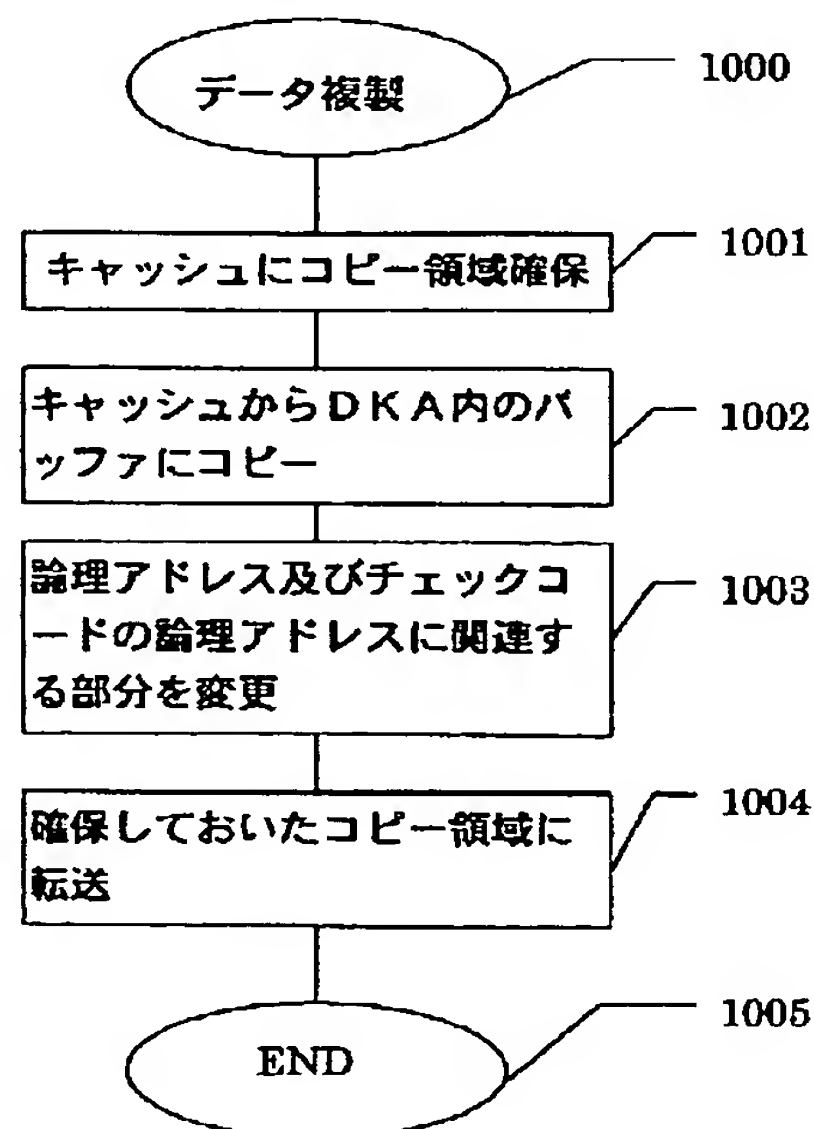
(54)【発明の名称】 ディスクサブシステム及びデータ複写方法。

(57)【要約】

【課題】論理ボリュームのデータを複製する際に、リモートコピー機能を使用するとホストチャネルを使用する為、処理装置からのアクセスに使用できるホストチャネルが減少する。

【解決手段】DKA内のDRRを使用してコピー元論理ボリュームのデータの判別情報(論理アドレス等)をコピー先ボリュームの判別情報に変更する(1003)事により、ホストチャネルを使用しない為、全てのホストチャネルが処理装置との接続に使用できる。

図 1 0



【特許請求の範囲】

【請求項 1】データを記録する記憶装置と、この記憶装置へのデータ転送を制御するディスクアダプタと、処理装置とデータ処理を行うチャネルアダプタと、前記記憶装置及び／または前記処理装置からのデータを一時的に記録するキャッシュメモリと、このキャッシュメモリに記録されたデータの判別情報を変更する手段とを備えたディスクサブシステム。

【請求項 2】前記判別情報は、論理アドレスである請求項 1 に記載のディスクサブシステム。

【請求項 3】前記判別情報を変更する手段として、前記記憶装置に記録するデータのパリティを作成する DRR を使用する請求項 1 に記載のディスクサブシステム。

【請求項 4】記憶手段及び／または処理手段と接続されこれらからのデータを一時的に記録するバッファ手段と、

このバッファ手段と前記処理手段との間のデータ転送を制御する手段と、

前記バッファ手段と前記記憶手段との間のデータ転送を制御する手段と、

前記記憶手段内で一対の論理ボリュームを正ボリュームと副ボリュームとして指定する手段と、

前記正ボリュームのデータを前記バッファ手段に読み出す手段と、

読み出した前記データの論理アドレスを前記副ボリュームの論理アドレスに変更する手段と、

前記論理アドレスを変更したデータを前記副ボリュームに転送する手段と、を備えたディスクサブシステム。

【請求項 5】記憶装置に記録したデータの複写方法であって、

処理装置からデータを複写するコマンドを受領するステップと、

前記データの複写先ボリュームの判別情報を設定するステップと、

キャッシュメモリ内に複写データを記録する領域を設定するステップと、

前記データをキャッシュメモリから読み出すステップと、

前記読み出したデータの論理アドレスを前記複写先の判別情報に変更するステップと、

論理アドレスを変更したデータを前記設定した領域に送るステップと、を有するデータ複写方法。

【請求項 6】前記判別情報は、論理アドレスである請求項 5 に記載のデータ複写方法。

【請求項 7】記憶装置に記録するデータのパリティを作成する DRR 内のバッファに前記データをキャッシュメモリから読み出し、前記 DRR にて前記判別情報を変更する請求項 5 に記載のデータ複写方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本願発明は、複数の論理ボリュームから構成されるストレージシステムにおいて、同一の制御装置内での論理ボリュームのペア及び複製作成に関する。

【0002】

【従来の技術】論理ボリュームの複製を作成する際には、まずペアとなるデータを作成し、その後にこのペアを分割する事で複製を作成する。このペア作成段階において単純なコピー動作を行うと、データはもちろんのこと、データの判別情報として使用される論理アドレスやチェックコードまでが同一となってしまう。そこでペアの新たな格納先を設定するためには、論理アドレスを新しく設定しなければならない。また、その後の入出力の基づき一度作成した複製を更新する際には、チェックコードの更新を行わなければならない。

【0003】これら新たなの設定を行うため、従来技術としては、特表平 8-509565 号公報に記載のような、いわゆるリモートコピー機能を使用し、制御装置があたかも処理装置であるかのように振る舞って、論理ボリュームのデータをホストインタフェースに送り出し、他のホストチャネル（CHA）にてこのデータを受信し、論理ボリュームに出力することにより、論理ボリュームのペアを作成した後に複製を作成していた。つまり、複製を作成する論理ボリュームのデータを一度ホストインタフェースに送り出し、新たなデータとしてホストチャネルで受信する事でペアの論理アドレス及びチェックコードの変更を行っていた。

【0004】しかし、従来技術では制御装置が処理装置として振る舞う為にホストチャネルを使用する。同一制御装置内での複製作成の場合にはホストチャネルを複数個使用し、その分、通常のホスト接続に使用できるチャネル数が減少する。

【0005】

【発明が解決しようとする課題】上記のリモートコピー機能を利用して論理ボリュームの複製を作成した場合には、複製データを転送する為にホストチャネルを使用する、これらのホストチャネルは処理装置との接続には使用できないという問題点が発生する。

【0006】本願発明の目的は、ホストチャネルを使用することなく論理ボリュームの複製を作成する事にある。

【0007】

【課題を解決するための手段】上記課題は、記憶手段及び／または処理手段と接続されこれらからのデータを一時的に記録するバッファ手段と、このバッファ手段と処理手段との間のデータ転送を制御する手段と、バッファ手段と記憶手段との間のデータ転送を制御する手段と、記憶手段内で一対の論理ボリュームを正ボリュームと副ボリュームとして指定する手段と、正ボリュームのデータをバッファ手段に読み出す手段と、読み出したデータ

の論理アドレスを副ボリュームの論理アドレスを変更する手段と、論理アドレスを変更したデータを副ボリュームに転送する手段とを備えるディスクサブシステムとすることにより達成できる。

【0008】

【発明の実施の形態】図1は本実施例のシステム構成を示す。制御装置102は、処理装置101からのデータ及び記憶装置103からのデータを一時的に保存するバッファ手段としてのキャッシュメモリ107と、処理装置101とのデータ転送を制御するあるチャンネルアダプタ（図ではCHAと記載）109、記憶装置101とのデータ転送を制御するディスクアダプタ（図ではDKAと記載）108、制御装置102の制御情報を記録する共有メモリ110とこれらを接続するバス120から構成され、磁気ディスク装置等の記憶媒体を備えた記憶装置103を制御し、処理装置101からの要求を実行する。

【0009】共有メモリ110に記録された制御情報は、共有メモリ110からチャンネルアダプタ109又はディスクアダプタ108にバスを経由して転送し、チャンネルアダプタ109及びディスクアダプタ108にあるプロセッサで使用する。尚、本願発明においてデータとは、ユーザデータ、ユーザデータの判別情報として使用するユーザデータの論理アドレス及びチェックコードからなる。更にチェックコードはユーザデータ及び論理アドレスをチェックするデータからなる。

【0010】CHA109及びDKA108は、その内部にバス120と接続するハードウェアであるバスアダプタ（図示せず）、データを転送を制御するプロセッサ1081（CHA109内は図示せず）等を備える。更にDKA108は記憶装置103記憶装置103に記録するデータのパリティ作成及び壊れたデータの回復をプロセッサ1081の指示に従って行うDRR（Data Recover and Reconstruct回路）1082を備えている。更にDRR1082内部にはパリティ作成・データ回復を行う際に利用するバッファ1083を備えている。尚、このバッファ1083はDRR1082内部に存在するとは限らない。

【0011】まず、ペア及び複製の作成・削除についてのコマンドセットを図2に示す。コマンドセットは、複製作成の為の正ボリュームと副ボリュームを指定してペアを作成するペア作成コマンド211と、正ボリュームと副ボリュームとのペアを解消するペア削除コマンド212と、副ボリュームに複製を作成するコマンド213と複製作成後に再度ペア有り複製未状態にする再同期コマンド214から成る。

【0012】次に、制御情報について説明する。図3にコピー元論理ボリューム（正ボリューム）とコピー先論理ボリューム（副ボリューム）とのペア情報300の詳細を示す。ペア情報300は、各ペア毎にペア番号30

1とペア状態302と正ボリューム番号（正ボリューム論理アドレス）303と副ボリューム番号（副ボリューム論理アドレス）304と正ボリュームの複製がどこまで終了しているのかを示す複製ポインタ305の情報を持つ。この情報は、ペア状態を遷移させる際の制御用の情報として共有メモリ110に記録する。

【0013】次にペア情報300のペア状態302の内容について図8を用いて説明する。図8にはペアの状態遷移を示す。ペア無し状態801とはペアを形成していない状態であり、ペア有り複製未状態802とはペアは存在するが副ボリュームに複製を生成していない状態である。

【0014】この複製を生成していない状態では、正副の全ボリュームコピーが終了したが、処理装置101からの入出力に同期して副ボリュームへの出力は実施されない。このため定期的に正ボリュームから副ボリュームへの差分データのコピーが実施されている状態である。

【0015】ペア有り複製中状態803とは副ボリュームに複製を生成中の状態、つまり正ボリュームのデータが副ボリュームにコピーされている状態である。この状態での正ボリュームへのライトは副ボリュームには反映されない。従って、正ボリュームへのライトが発生すると、正ボリュームと副ボリュームとの内容は不一致になる。つまり、副ボリュームには複製作成コマンド213が発行されたときの正ボリュームの複製が作成される。

【0016】ペア有り複製有り状態804とは副ボリュームに複製の生成を完了した状態である。この段階において副ボリュームはリード／ライトの何れも可能となり正ボリュームとは独立して作業を行う事が出来る。

【0017】複製作成コマンド213が発行された後に正ボリュームにライトがあった場合には、ライト位置は図4に示す差分情報として記録しておき、再同期コマンド214が発行され、ペア有り複製未状態802を作成する際に使用される。

【0018】ペア削除コマンド212によりどの状態からでもペア無し状態801に遷移する。

【0019】ペア作成コマンド211により、ペア無し状態801からペア有り複製未状態802に遷移する。複製作成コマンド213によりペア有り複製未状態802からペア有り複製中状態803に遷移する。複製全体処理500が終了するとペア有り複製中状態803からペア有り複製有り状態804に遷移する。再同期コマンド214によりペア有り複製有り状態804からペア有り複製未状態802に遷移する。

【0020】図4に差分情報401を示す。差分情報401は、ペア番号301を引数とする配列になっており、各エントリは差分ビットの集合から成る。差分ビットは対応するデータについて正ボリュームと副ボリュームの間に差異が無い場合には0を設定し、差異が存在する場合に1を設定する。この情報は、制御用の情報とし

て共用メモリ110に記録する。この差分情報401は一度複製を作成した後に再び正ボリュームと副ボリュームとの整合をとる（再同期させる）際に使用する。

【0021】この差分情報401は、処理装置101からデータのライト要求があった場合には、対応する差分ビットを1に設定する。また、複製処理を実施することによりデータが一致した場合には、対応する差分ビットを0に設定する。

【0022】図7にコマンドブロックの構造を示す。コマンドブロック700はコマンドコード701と正ボリューム番号702と副ボリューム番号703から成る。

【0023】制御装置102は処理装置101からコマンドブロック700を受領するとコマンドブロック700の正ボリューム番号702及び副ボリューム番号703と一致する正ボリューム番号303及び副ボリューム番号304をもつエントリをペア情報300から検索する。このエントリからペア状態302を取得し、このペア状態302とコマンドブロックのコマンド701から図2に示す処理を行う。

【0024】ペア作成コマンド211が指定するペアの状態がペア無し221の場合、新規ペア作成231を実施する。新規ペア作成231では、まずペア情報300を元に処理装置101から指定されたボリュームに空きエントリを見つけて副ボリューム用のエリアとして割り当てる。そしてこの割り当てに基づきペア作成コマンド211で指定した正ボリューム番号303と副ボリューム番号304とをそれぞれペア情報300の正ボリューム番号303のエリアと副ボリューム番号のエリアに設定し、複製ポインタ305を初期化し、複製全体処理500をコールする。

【0025】この新規ペア作成の際に、RAIDレベルが異なる論理ボリュームを副ボリュームのエリアとして指定する事で、同一のデータに異なるアクセス特性を持たせる事が出来る。尚、この新規ペア作成の段階では、副ボリューム用のエリアを確保するだけであり、データのコピーは未だ行われておらず、副ボリュームの格納先が決められるのみである。

【0026】ペア作成コマンド211を受けた時のペア状態がペア無し221以外の場合、原則として制御装置102は何も行わない。ペアが有るにもかかわらず再度ペアを作成するのは、正ボリュームが複数の異なる論理ボリュームを副ボリュームとして定義し、それぞれ異なるペアと定義して複製を作成する場合である。異なるペアと定義して複製を作成する場合とは、例えば処理装置101から複製指示コマンドを受け付けた時刻を指定時刻とし、一定時間毎に次々に複製を作成する事により、正ボリュームの履歴を作成したい場合等が考えられる。

【0027】また、ペア削除コマンド212を受けた時のペア状態がペア無し221の場合、制御装置102は何も行わない。この場合、ペアが存在していないのでこ

のペア削除コマンド212はリジェクトされる。

【0028】ペア削除コマンド212を受けた時のペア状態がペア有222, 223の場合、制御装置102はペア情報300の対応とするペアのエントリを初期化する。

【0029】複製形成コマンド213を受けた時のペア状態がペア無し221の場合、制御装置102は何も行わない。

【0030】複製形成コマンド213を受けた時のペア状態がペア有複製未222の場合、差分情報401をボリューム全領域に対応する差分ビットを1に設定し、複製ポインタ304を初期化し、複製全体処理500をコールする。

【0031】複製形成コマンド213を受けた時のペア状態がペア有複製有223の場合、制御装置102は何も行わない。

【0032】再同期コマンド214を受けた時のペア状態がペア無し221の場合、制御装置102は何も行わない。この場合、ペアが存在していないのでこのペア削除コマンド212はリジェクトされる。

【0033】再同期コマンド214を受けた時のペア状態がペア有り複製未222の場合、制御装置102は何も行わない。

【0034】再同期コマンド214を受けた時のペア状態がペア有り複製有り223の場合、ペア状態をペア有り複製未222に遷移する。

【0035】次に複製処理について、図5と図6を用いて詳細に説明する。

【0036】まず、複製全体処理500は、ペア作成コマンド211又は複製形成コマンド213を受領時にコールされる。コール時の入力情報はペア番号を指定する。複製を行いたいデータがキャッシュ107上に存在していない場合には、記憶装置103から読み出す。

【0037】最初にステップ501で入力情報のペア番号と一致するペア番号のエントリをペア情報300から検索し、当該エントリのペア状態300を取得し、ペア状態302がペア有り複製未状態802であることを確認する。ペア有り複製未状態802でなければ、複製処理を実行できないと判断し処理を終了する。

【0038】ペア有り複製未状態802であれば、単位複製処理600をペア番号を入力情報としてコールする502。

【0039】図6に単位複製処理600の詳細を示す。

【0040】単位複製処理600がコールされると、ステップ601において入力情報のペア番号と一致するペア番号のエントリをペア情報300から検索し、当該エントリの複製ポインタ305を取得する。次にステップ602においてステップ601で取得した複製ポインタ305が示すアドレスから差分ビットが1になっているデータを差分情報401から検索する。ステップ603

ではステップ602で検索したデータを正ボリュームからキャッシュ107に読み出す。ステップ604ではステップ603で読み出したデータをキャッシュに別の領域を取得し、データのコピーとデータに付与されている論理ボリューム番号を正ボリューム番号303から副ボリューム番号304へ変更する。ステップ605ではステップ604でコピーしたデータを副ボリュームへ出力する。

【0041】ステップ609では副ボリュームへ出力したデータ分の差分情報401を0に設定する。

【0042】ステップ606では副ボリュームへ出力したデータ分だけ複製ポインタ305を進める。ステップ607ではステップ606で変更した複製ポインタ305をペア情報300に記録し、単位複製処理を終了しコール元へリターン608する。

【0043】ここでペア無し状態801からペア作成コマンド211にてペアを作成し、複製全体処理500がコールされて更に単位複製処理600がコールされた場合のステップ603及び604について更に説明する。

【0044】先にも述べたように、最初にペアを作成しデータを複写する際に論理アドレスは新たに設定する必要がある。また、この論理アドレスの設定或いは変更に伴うチェックコードの変更も必要となる。本願発明ではこの新たな論理アドレスの設定をDKA108内のDRR1082を用いて行っている。DRR1082は本来、RAID構成を採るディスクサブシステムにおいて、データからパリティを作成、並びにパリティ及びその他のデータから障害データの回復を行うためのものである。

【0045】このパリティ作成の際にパリティに論理アドレスを割り振る機能を利用して本願発明では副ボリュームに作成する複製の論理アドレスを設定する。このため、本願発明における論理ボリュームの複製手続きはキャッシュ107とDKA108とによって行われている。尚、キャッシュ107での複製をするにあたり、論理アドレスの変更を行う機能及びバッファを備えた構成要素を別個に備えるのであれば、DKA108以外の要素、場所との組み合わせでコピーを行う事が出来るのはもちろんの事である。

【0046】図10を用いてステップ604を説明する。ステップ1000において、制御装置102がキャッシュ107のデータの複製をDRR1082のハードウェアに指示を出す。そしてステップ1001において、制御装置102はキャッシュ107上に複製を置くための領域を確保する。次にステップ1002においてDKA108のプロセッサ1081は、ステップ603においてキャッシュ107に読み出された対象データ(=正ボリュームのデータ)をDRR1082内部のバッファ1083に読み出す。

【0047】ステップ1003において、DRR108

2を使用する事により読み出したデータの論理アドレスを新規ペア作成231が発行された際に設定されたボリューム番号に変更する。その際に、変換されたボリューム番号に併せてチェックコードも変更する。最後にステップ1004において、この論理アドレスを変更したデータを、やはり新規ペア作成231で設定されたキャッシュ107内の副ボリューム用エリアに転送する。以上によりキャッシュ上でのコピーを行う。

【0048】尚、通常、キャッシュ107はデータを管理する際にそのディレクトリにデータの論理アドレス等を記憶しておき、データの転送の際にはディレクトリに記録されたアドレスを使用する。そのため、キャッシュ107上でデータの論理アドレスを変更してもこのディレクトリのアドレスを変更しない限り副ボリュームへの出力を行う事が出来ない。よって本実施例においては上記のようにDRR1082内のバッファ1083を用いてキャッシュ107上でのコピーを作成した。しかし、データの論理アドレスとともに、キャッシュディレクトリのアドレスも変更する手段を備えれば、DRR1082内のバッファ1083を使用しキャッシュ107上にコピーを作成することなく、キャッシュ107上のデータ及びディレクトリの論理アドレスを変更した後にそのデータを副ボリュームに出力する事で複製を作成してもよい。

【0049】単位複製処理600からリターン後、複製完了判定503を行う。つまり、複製ポインタ305が論理ボリュームの末尾アドレスと一致していない場合には複製未完了と判定し、ステップ501から繰り返す。一方、一致している場合は複製完了を判定し、ステップ504に進む。

【0050】ステップ504では、複写終了したデータを、キャッシュ107から副ボリュームに転送されることで、複写が指示された時点での正ボリュームのデータの副ボリュームへの複写が完了する。そしてペア状態302を複製有りに変更し、複製処理を終了する。

【0051】次に、複製処理中に正ボリュームのデータへの変更要求が処理装置101から制御装置102にあった場合の処理について、図9を用いて説明する。図9には、複製処理中に正ボリュームのデータへの変更要求があった場合の処理(先行コピー)のフローチャートを示す。

【0052】ステップ901にて入力情報のボリューム番号と同じ正ボリューム番号をもつペア情報300のエントリを検索し、ペア番号を求める。

【0053】ステップ902にてステップ901で求めたペア番号のエントリからペア状態302を求め、ペア有り複製中状態803であるかどうか判定し、ペア有り複製中状態803でない場合には、ステップ909にて通常の出力処理を実行し、ペア有り複製中状態803である場合には、ステップ903に進む。

【0054】ステップ903では、ステップ901で求めたペア番号から差分情報400の対応するエントリを求め、処理装置101から変更要求のあったデータについて既に差分ビットが1であるかどうか判定し、差分ビットが1でない場合には、ステップ909にて通常の実出力処理を実行し、差分ビットが1である場合には、ステップ904に進む。

【0055】ステップ904では、処理装置101から変更要求のあったデータを正ボリュームからキャッシュ107に読み出す。

【0056】ステップ905では、ステップ904でキャッシュ107上に読み出したデータからキャッシュ107上に副ボリューム用のデータを複製する（先に示した図6のステップ604と同じ）。

【0057】ステップ906では、ステップ905で複製したデータを副ボリュームに出力する。

【0058】ステップ907では、上記ステップで副ボリュームに出力したデータの差分ビットから0に設定する。

【0059】ステップ909では、処理装置101から変更要求のあったデータを処理装置101から受け取り、正ボリュームに出力する通常の実出力処理を実行する。

【0060】以上のように、本願発明は、正ボリュームとペアとなるボリュームを指定して、ペアの作成を指示すると、正ボリュームの複製が副ボリュームとして作成される。この状態（図2における複製作成コマンド213が発行される前）では処理装置から副ボリュームへの書き込みは禁止されている。ペアの分割が指示されると複製作成コマンド213が実行され、正副各々のボリュームがホストから独立に運用可能となる。再同期が指示されると再同期コマンド214が実行され、副ボリュームに対する書き込みが禁止されるとともに、正ボリユー

ムと副ボリュームとの差分が図4に示す差分情報を元に副ボリュームにコピーされ、再び副ボリュームは正ボリュームの複製となり、再度分割可能な状態となる。

【0061】また、上記説明したように、本願発明ではDKA内部のDRRを正ボリュームの複製作成時に使用する事で、CHAを利用することなく正ボリュームの複製を同一ディスクサブシステム内の別の副ボリュームに作成することが出来る。

【0062】

【発明の効果】本発明により、ホストチャネルたるCHAを利用すること無く、指定した論理ボリュームの複製を作成することが出来る。これにより、複製作成に伴う制御装置の負荷の削減を図ることが出来る。

【図面の簡単な説明】

【図1】システム構成を示す図である。

【図2】図1のコマンド動作内容を示す図である。

【図3】ペア情報を示す図である。

【図4】差分情報を示す図である。

【図5】複製処理を示すフローチャートである。

【図6】単位複製処理順を示すフローチャートである。

【図7】コマンドブロックフォーマットを示す図である。

【図8】ペア状態遷移を示す図である。

【図9】複製処理中にデータ変更要求が発生した場合の処理順を示すフローチャートである。

【図10】判別情報変更処理を示すフローチャートである。

【符号の説明】

101…処理装置、102…制御装置、103…記憶装置、107…キャッシュメモリ、108…ディスクアダプタ、109…チャネルアダプタ、110…共有メモリ、1081…プロセッサ、1082…DRR、1083…バッファ。

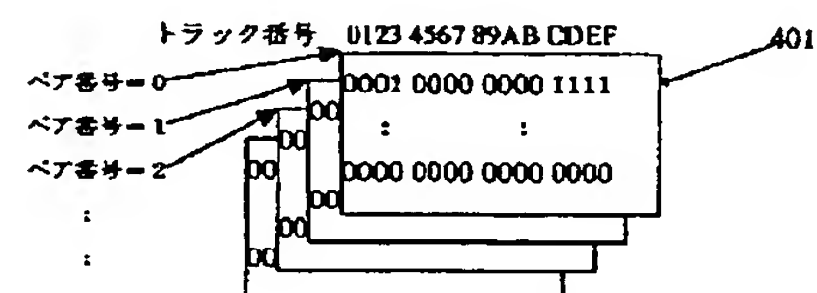
【図2】

図2

コマンド名称 (210)	ペア状態 (220)		
	ペア無し (221)	ペア有り未複製 (222)	ペア有り複製有り (223)
ペア作成 (211)	新規ペア作成 (231)	なし (232)	なし (233)
ペア削除 (212)	なし (241)	ペア削除 (242)	ペア削除 (243)
複製作成 (213)	なし (251)	複製開始 (252)	なし (253)
再同期 (214)	なし (261)	なし (262)	再同期開始 (263)

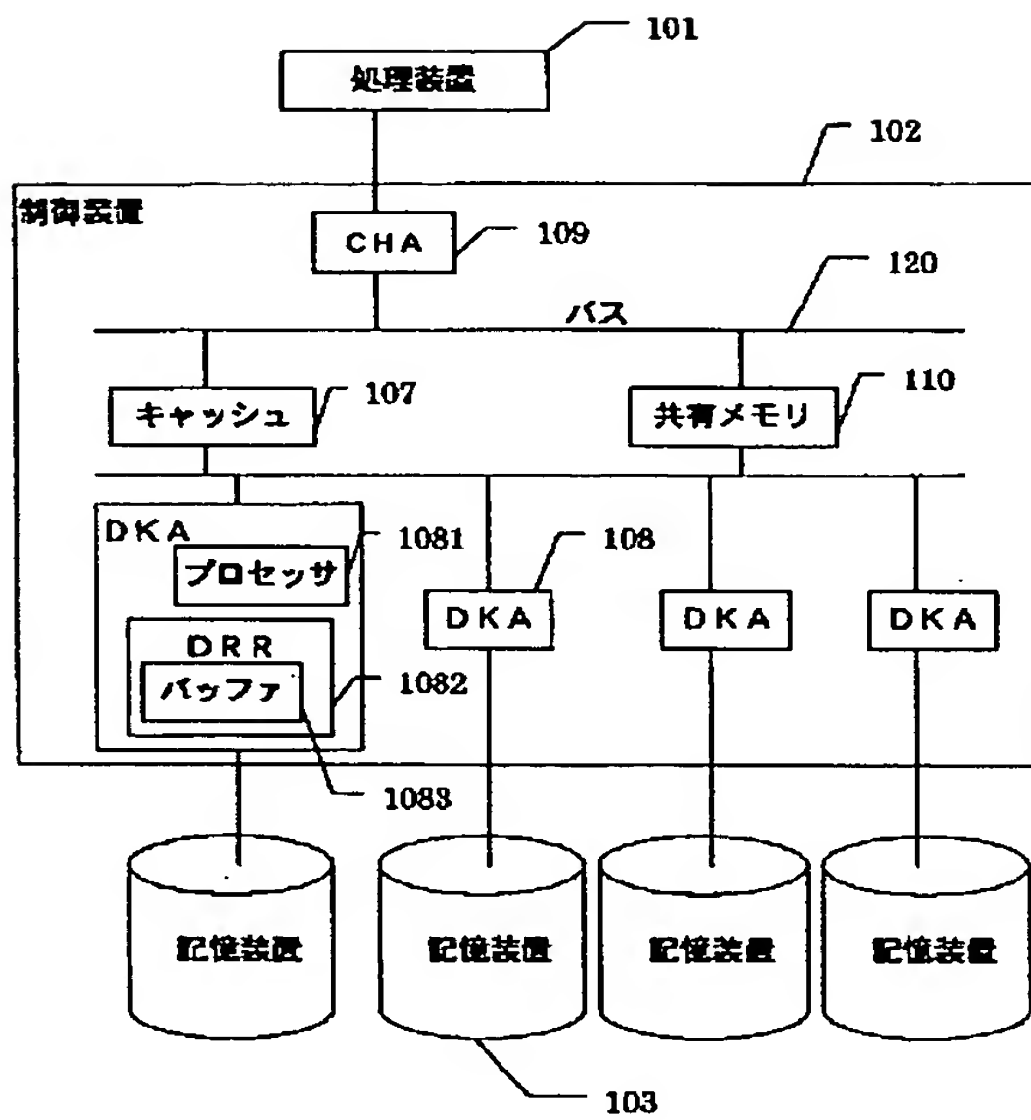
【図4】

図4



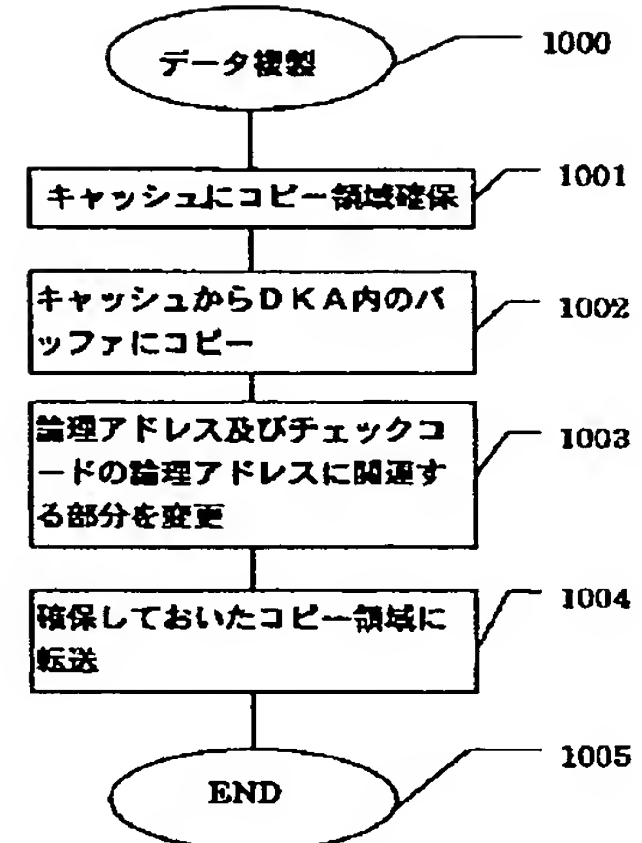
【図1】

図1



【図10】

図10



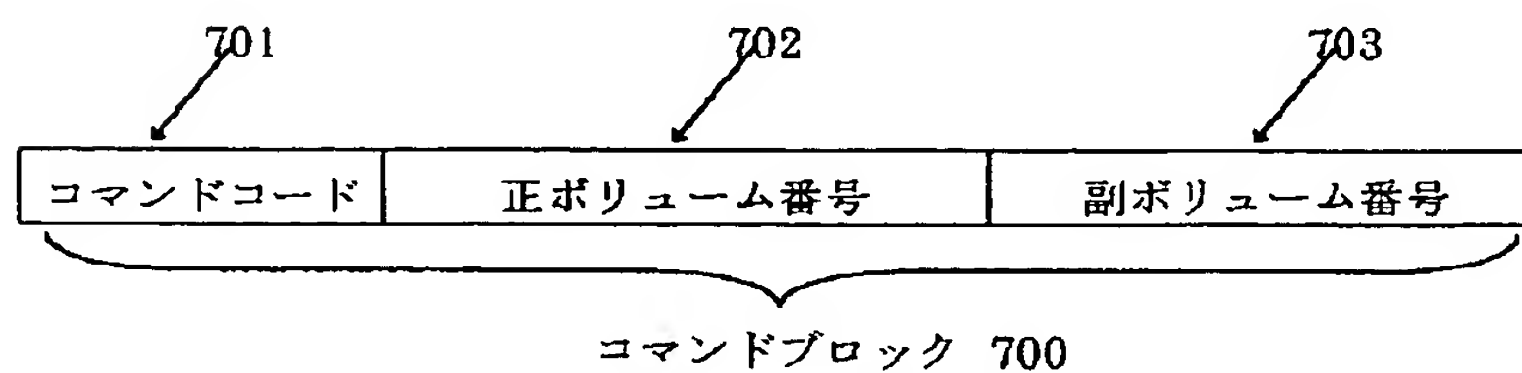
【図3】

図3

ペア番号	ペア状態	正ボリューム番号	副ボリューム番号	複製ポインタ
1	ペア有 複製未	1 2	6 4	0
:	:	:	:	:

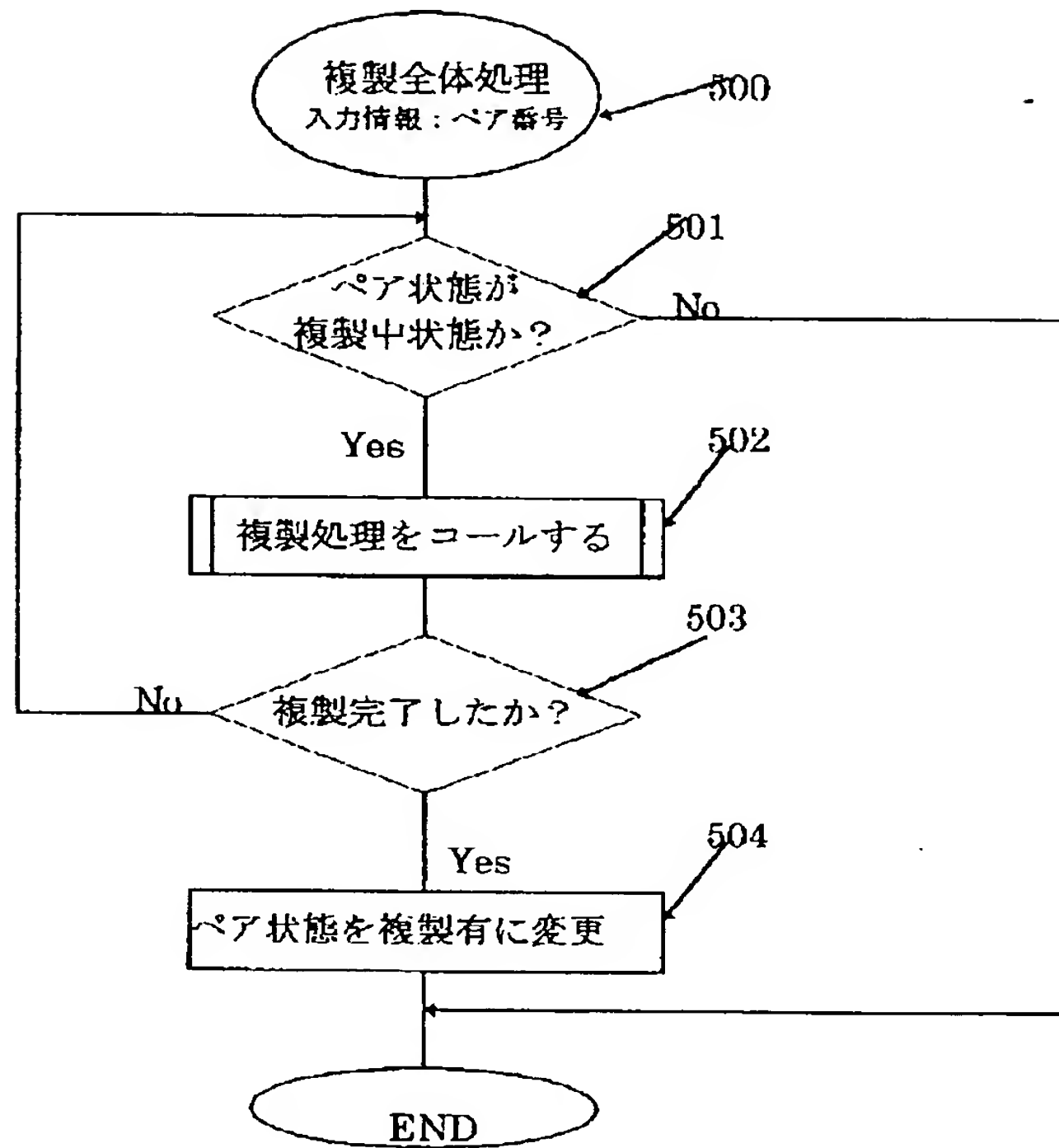
【図7】

図7



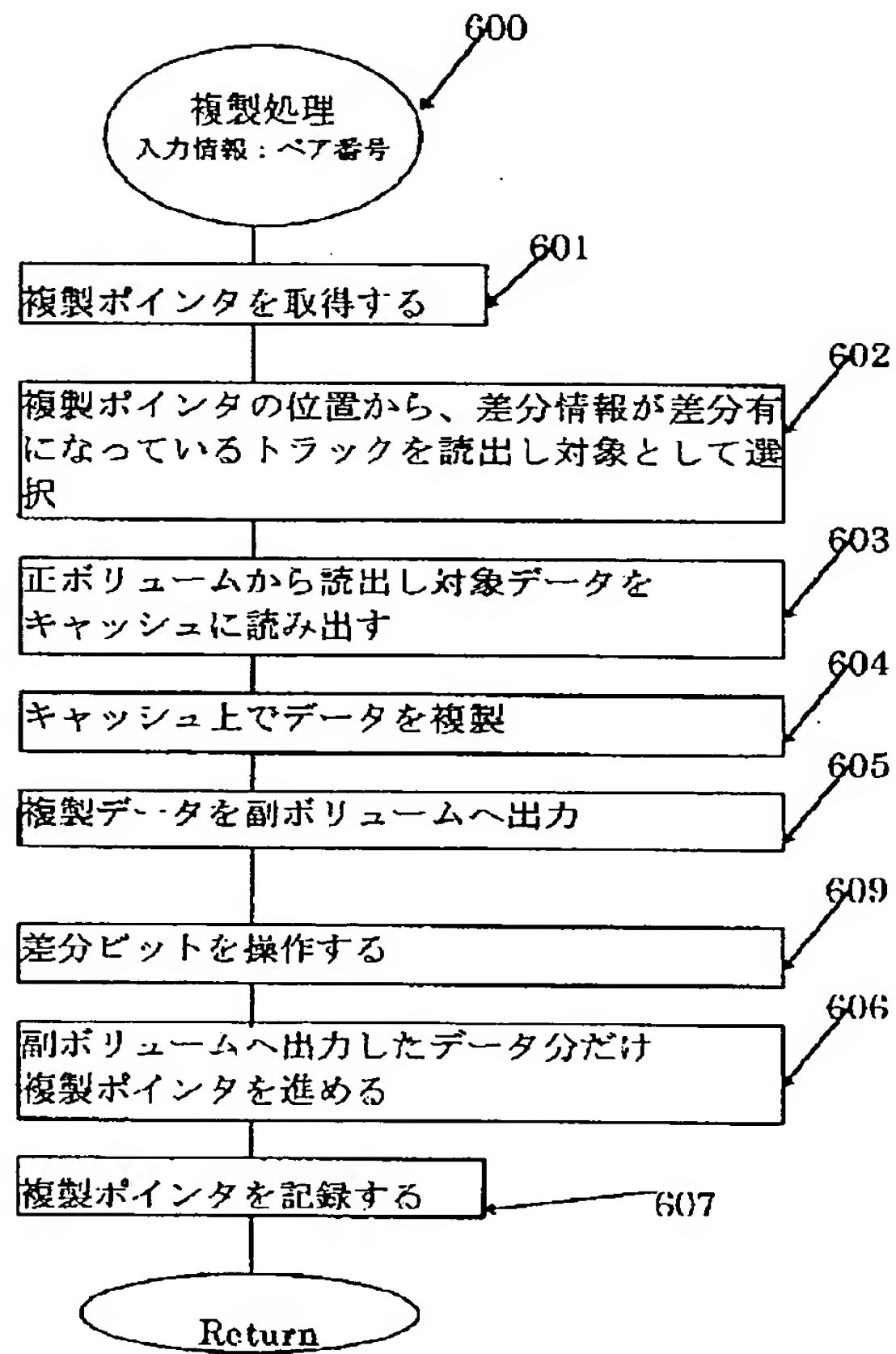
【図5】

図5



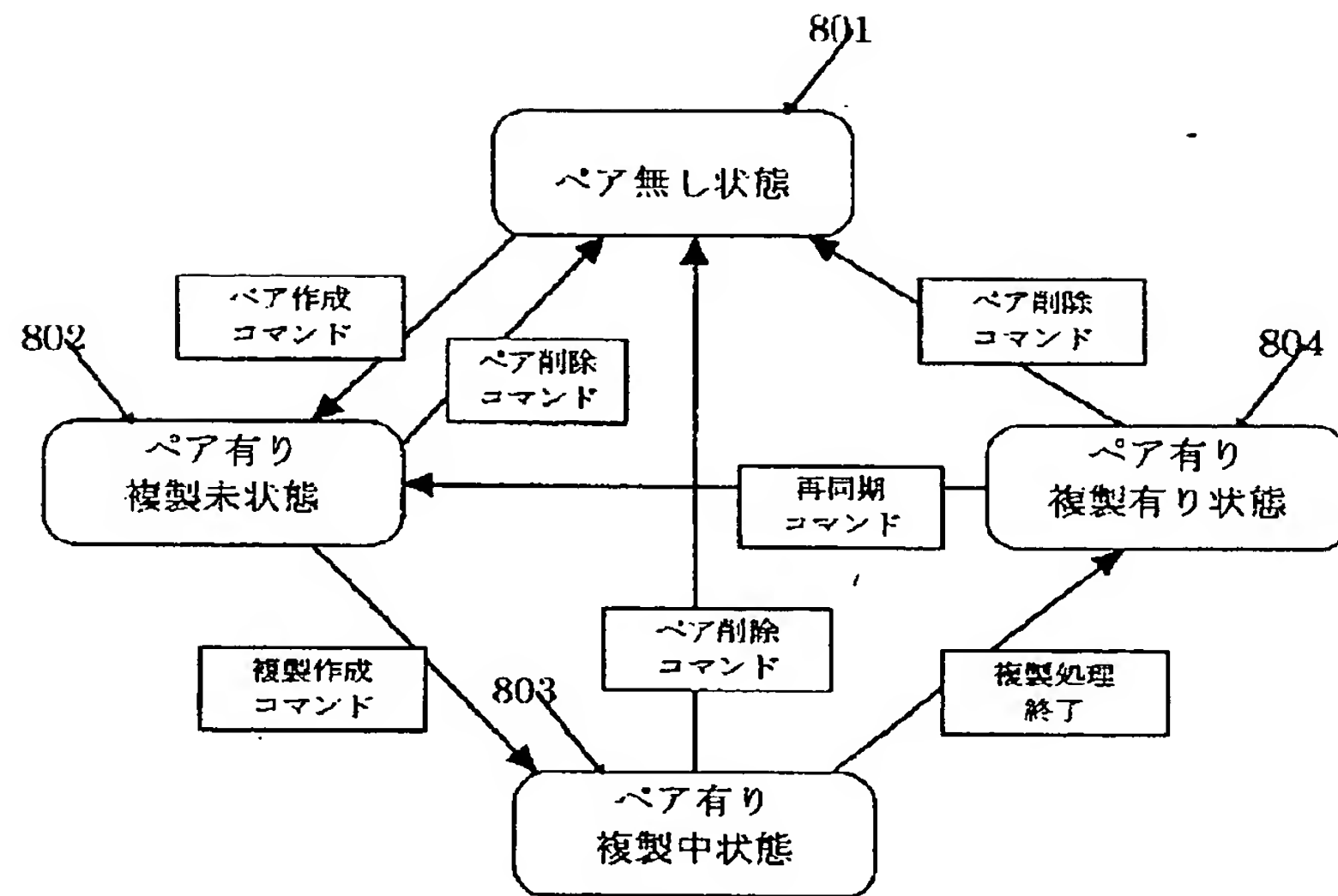
【図6】

図6



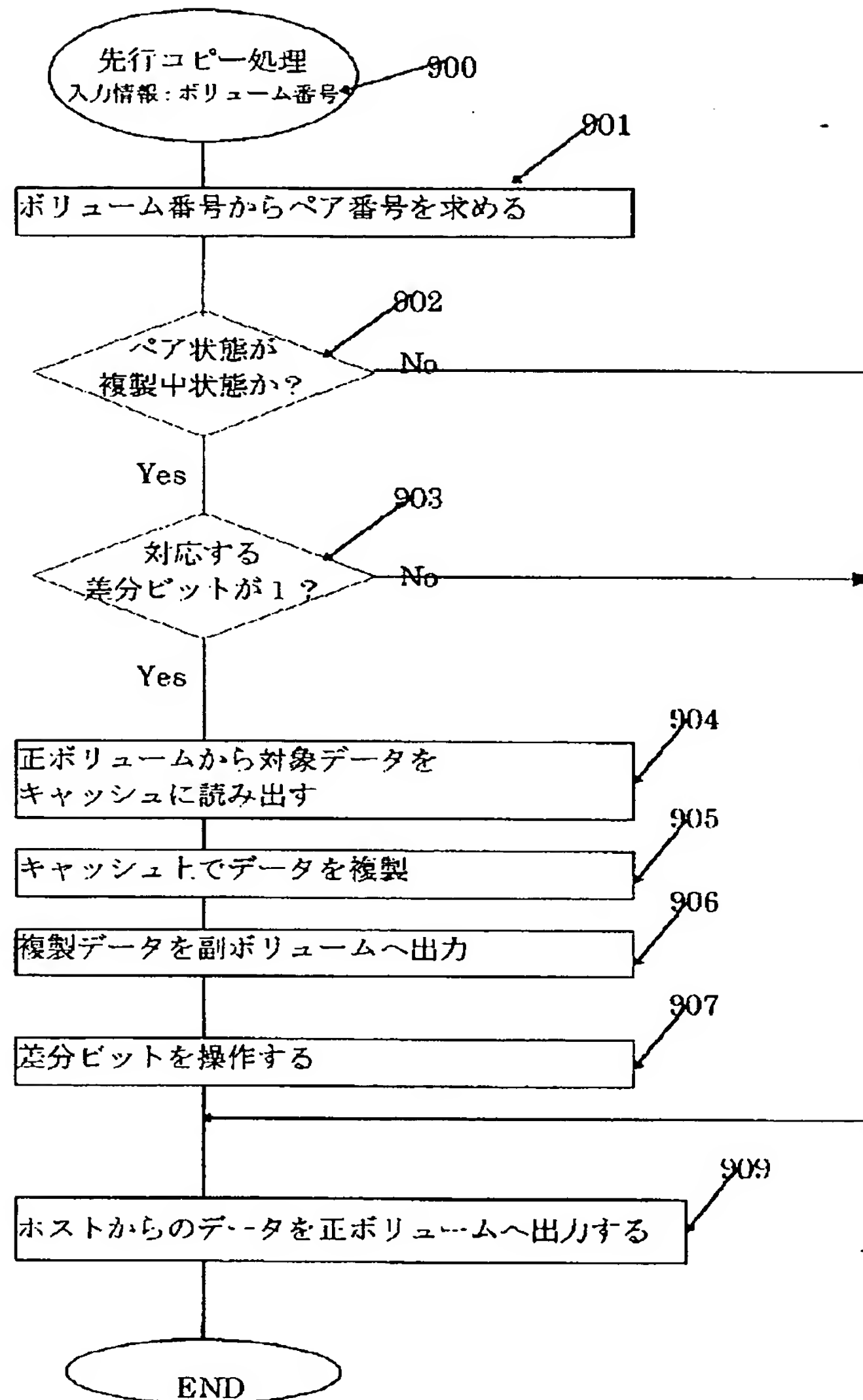
【図8】

図8



【図9】

図9



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.